A mountain ridge model for quantifying oblique mountain wave propagation and distribution

Sebastian Rhode¹, Peter Preusse¹, Manfred Ern¹, Jörn Unger mann¹, Lukas Krasauskas¹, Julio Bacmeister², and Martin Riese¹

¹Institute of Energy and Climate Research, Stratosphere (IEK-7), Forschungszentrum Jülich, Jülich, Germany
²National Center for Atmospheric Research (NCAR), Boulder, CO, USA

Correspondence: Sebastian Rhode (s.rhode@fz-juelich.de)

Abstract. Following the current understanding of gravity waves (GWs) and especially mountain waves (MWs), they have high potential of horizontal propagation from their source. This horizontal propagation and therefore the transport of energy is usually not well represented in MW parameterizations of numerical weather prediction and general circulation models. The lack thereof possibly leads to shortcomings in the model’s prediction as e.g. the cold pole bias in the Southern Hemisphere and the polar vortex breaking down too late. In this study we present a mountain wave model (MWM) for quantification of the horizontal propagation of orographic gravity waves. This model determines MW source location and associates their parameters from a fit of idealized Gaussian shaped mountains to topography data. Propagation and refraction of these MWs in the atmosphere is modeled using the ray-tracer GROGRAT. Ray-tracing each MW individually allows for an estimation of momentum transport due to both vertical and horizontal propagation. This study presents the MWM itself and gives validations of MW induced temperature perturbations to ECMWF IFS numerical weather prediction data and estimations of GWMF compared to HIRDLS satellite observations. The MWM is capable of reproducing the general features and amplitudes of both of these data sets and, in addition, is used to explain some observational features by investigating MW parameters along their trajectories.

1 Introduction

Gravity waves (GWs) are atmospheric waves for which gravity or buoyancy acts as the restoring force (Fritts, 1984). They are an important dynamical process and interact with the large scale flow and contribute to the generation of clouds. Since they propagate vertically through the atmosphere, they transport momentum from the lower atmosphere, or even the ground, to higher levels such as the stratosphere, mesosphere and lower thermosphere. This relocation of momentum is one of the drivers of the Brewer-Dobson-Circulation (BDC) and the main driver of the upper, mesospheric branch of the residual circulation (e.g. McIntyre, 1998; Fritts and Alexander, 2003). Various studies also argue for a significant role of gravity waves in the occurrence of Sudden Stratospheric Warming (SSW) events (e.g. Whiteway et al., 1997; Kidston et al., 2015) and even their shape (Albers and Birner, 2014; Ern et al., 2016; Song et al., 2020).
In addition, effects of GWs are a major uncertainty in climate projection and numerical weather forecast (Shepherd, 2014). On the one hand, GWs have a significant impact on the dynamics of the atmosphere and even larger scale climate phenomena, on the other hand, they have to be parameterized within general circulation models (GCMs). While larger scale GWs are resolved by the models, small scale GWs caused by the sub-grid-scale orography and convection are approximated by a parameterization scheme (e.g. Lott and Miller, 1997; Kim et al., 2003; Xie et al., 2020). One particular shortcoming of these parameterizations employed in GCMs is the commonly used column-wise calculation approach, which does not allow for the GW’s momentum to be transported horizontally, whereby the corresponding GW drag will be exerted above the mountains. However, a high potential of horizontal GW propagation has been found both in observations and model studies (e.g. Preusse et al., 2002; Sato et al., 2012; Krisch et al., 2017; Ehard et al., 2017; Strube et al., 2021).

Another shortcoming of current GCMs is poor performance in the Southern Hemisphere dynamics, leading to the polar vortex breaking down too late in spring and too low temperatures above the South Pole (cold-pole bias) (Butchart et al., 2011). Studies have shown that introducing north-south hemispheric asymmetry in orographic GW drag tuning can alleviate this shortcoming significantly (Garcia et al., 2017). Enhancing the Southern Hemisphere GW drag by a factor of 2 over that in the Northern Hemisphere leads to much better agreement between the simulated dynamical and chemical structure of the Southern Hemisphere polar vortex and observations while optimizing the model’s performance in the Northern Hemisphere. In a different study, (Polichtchouk et al., 2018) show that an non-orographic GW parameterization with increased drag improves the models prediction of the polar vortex breakdown. A possible reason for the underestimation of gravity wave drag in the southern polar vortex (McLandress et al., 2012) is the missing representation of oblique GW propagation (Wu and Eckermann, 2008; Strube et al., 2021; Perrett et al., 2021) and possibly the under-representation of small scale Ocean Islands (Alexander et al., 2009, 2010; Garfinkel and Oman, 2018; Hindley et al., 2021). Both processes could play an important role in explaining the missing wave drag at 60°S in climate and chemistry models.

Although GWs can be excited by various processes, e.g. convection, jet imbalances and even volcanic eruptions (Wright et al., 2022; Ern et al., 2022), one of the major and most predictable sources of gravity waves is wind flow over orography, by which air parcels are displaced vertically. These stationary (with respect to the ground) mountain waves (MWs) propagate through the atmosphere, both vertically and horizontally. In the middle atmosphere they can be measured from satellites (e.g. Eckermann and Preusse, 1999; Preusse et al., 2002; Jiang et al., 2004; Eckermann and Wu, 2012). The strength of the excited MWs depends on the height and shape of the orography, velocity of the low level flow and the propagation conditions due to the winds above. At midlatitudes, westerly winds prevail in the troposphere and north-south oriented ridges are expected to be particularly effective MW sources. Accordingly, the Rocky Mountains and Andes are regions in which particularly high MW activity is expected. Indeed, for both mountain chains severe clear air turbulence due to mountain waves has caused major aviation incidents (e.g. Smithsonian Magazine, 2005; Boldmethod, 2016; Aviación Global, 2019). However, climatologies of GWs for the mid-stratosphere show only moderate GW activity over both the Rocky mountains and the Andes north of 40°S (e.g. Geller et al., 2013; Ern et al., 2018; Hindley et al., 2020). Likewise, the highest mountains on Earth, the Himalayas, have only moderate impact on middle atmosphere GW distributions. In order to understand this low stratospheric GW activity, the observations need to be combined with a model study focusing on mountain waves.
In this study, we present a mountain wave model (MWM) capable of estimating the sources of orographic gravity waves similar to the approach by Bacmeister (1993) and Bacmeister et al. (1994). The propagation of the so-determined MWs throughout the atmosphere is modeled using the ray-tracer GROGRAT (Marks and Eckermann, 1995). Refraction of GWs due to horizontal gradients and time dependence of the background fields is considered within the ray-tracer. Results include 4-dimensional momentum flux distributions, drag exerted on the background winds and estimations of the residual temperature perturbations caused by the MWs. Compared to previous studies, we aim for a higher accuracy in terms of mountain source detection by fit of an idealized mountain shape at arbitrary angles. Additionally, we are considering wind blocking and surface friction effects in the low level winds to improve modeled MW amplitudes and field characteristics. We are validating our model against satellite and model data to higher degree of detail and emphasis on the MWM than previous studies (e.g. Jiang et al., 2002, 2004). For this, we consider data at altitudes of 15-25 km, which is low enough for a comparison of the effect of primary MWs before and after wind filtering in the atmosphere. Satellite data at these low altitudes have not been used for such a comparison before.

The described MWM is used to explain GW features in satellite data by investigating their wave characteristics from source to observation altitude. MW propagation patterns throughout the year are predicted and agree to previous studies of oblique MW spread. Therefore this model might be used for identifying MW propagation pattern in further studies for improving MW parameterizations in GCMs by approximating their horizontal spread.

This article is organized as follows: First, Sect. 2 introduces the data used in this study. A general description of the mountain wave model is given in Sect. 3, which describes the detection of mountain wave sources, estimation of launch parameters and modeling of the propagation. In addition, the post-processing of MWM data resulting in reconstructions of residual temperature and gravity wave momentum flux (GWMF) distributions is discussed. Following this, a brief validation of the model is given in Sect. 4 including an investigation of the detected scales and a comparison to ECMWF operational analysis temperature data. In Sect. 5 the model’s capability to predict horizontal GW propagation is shown by comparison with HIRDLS satellite data. Predicted GW parameters and their change through the observation altitudes an wind blocking are considered as possible causes of some observational features. In addition, predictions of horizontal GWMF pattern throughout the year are shown, which give first insight in the universality of horizontal propagation. Finally, the results are summarized in Sect. 6 and concluding remarks are given.

2 Data

This study uses topography data for the ridge finding, atmospheric background winds and temperature for MW propagation modelling. Satellite observations of gravity wave momentum fluxes (GWMFs) are used for a validation and comparison. The used data sets are presented in the following sections.

2.1 Topography data

The underlying topographic elevation data used in this study is taken from the ETOPO1 GLOBAL RELIEF MODEL (Amante and Eakins, 2009; Center, 2009) data set. This data is available in two versions: one describes the bedrock elevation only, the
other considers also the ice surface, i.e. glaciers and ice sheets. Since we are interested in the elevation encountered by the low level flow, we use the data including the ice surface. This data set models the earth’s surface, including ocean bathymetry, on an 1 arc-minute resolution and is combined from multiple global and regional data sets. However, we set all regions below sea level, which are not relevant for our analysis, to zero to approximate an ocean surface.

2.2 Atmospheric backgrounds

The atmospheric background wind and temperature data used for ray-tracing of MWs are generated from ECMWF ERA5 reanalysis data (Hersbach et al., 2020; C3S, 2017), sampled on a $0.3^\circ \times 0.3^\circ$ grid. For our ray-tracing experiments, we want the background to contain all global and synoptic scale features but no GWs, which are potentially resolved by the numerical weather prediction (NWP) model. Following Strube et al. (2020), a scale separation approach is therefore applied to the data set using a zonal Fourier transform with a cutoff zonal wave-number of 18. In the meridional direction the data is smoothed using a 3rd order Savitzky-Golay filter of 31 subsequent data points ($\sim 9^\circ$ total window width). For the use in GROGRAT, the smoothed background is sampled onto a grid of $2^\circ$ latitude and $2.5^\circ$ longitude. In the vertical the data is interpolated to equidistant altitudes of 0.5 km spacing.

2.3 HIRDLS satellite data

The horizontal GWMF distributions generated by our MWM for the year 2006 are validated and compared in Sect. 5.1 to satellite data from the HIgh Resolution Dynamics Limb Sounder (HIRDLS, Gille et al. (2003)) instrument. Here we use a data set specially prepared for the UTLS region, spanning the altitudes from 14 km to 25 km. As discussed in Strube et al. (2020), below 20 km altitude zonal wave-numbers of 10 and higher need to be taken into account in order to describe the background. These cannot be self-consistently estimated from a single-track low Earth orbit satellite (Salby and Callaghan, 1997). In order to isolate the small scale gravity wave (GW) contributions, ERA5 background data with an altitude-dependent zonal wave number cutoff has been subtracted from the retrieved HIRDLS temperature measurements. Below 10 km altitude a cutoff at zonal wave number 20, above 20 km a cutoff at 6 has been used. In between, the cutoff decreases linearly from 20 to 6. The resulting vertical profiles of temperature residuals have been used for the calculation of GWMF as described in Ern et al. (2018). In order to analyze the lower stratosphere (20 km and below) and simultaneously avoid influence of the tropopause, the vertical window of the MEM/HA method (Preusse et al., 2002) was reduced from 10 km, which is usually used for stratospheric altitudes (e.g. Ern et al., 2004), to 5 km. Such a reduced window size is adequate, since in the lower stratosphere the average vertical wavelengths are much lower than in the mid stratosphere and mesosphere (e.g. Chane-Ming et al., 2000; Yan et al., 2010; Ern et al., 2018). In addition, the HIRDLS data set has been high-pass filtered in terms of vertical wavenumbers using a 5th order Butterworth filter with vertical wavelength cutoff at 12 km, similar to Ehard et al. (2015).

For this paper, GWMF is binned within rectangular overlapping bins of $15^\circ$ in longitude and $5^\circ$ in latitude sampled every $5^\circ$ in longitude and $2.5^\circ$ in latitude. The vertical resolution of the data set is 1 km, which corresponds to the vertical resolution of HIRDLS. Note that due to the vertical window of 5 km used in the MEM/HA method the given levels are representative for $\pm 2.5$ km around the corresponding altitude.
Table 1. Band-pass intervals used for detecting mountain ridges in the topography for this study.

<table>
<thead>
<tr>
<th>Interval No.</th>
<th>lower boundary [km]</th>
<th>upper boundary [km]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>80</td>
<td>150</td>
</tr>
<tr>
<td>2</td>
<td>150</td>
<td>250</td>
</tr>
<tr>
<td>3</td>
<td>250</td>
<td>400</td>
</tr>
<tr>
<td>4</td>
<td>400</td>
<td>600</td>
</tr>
<tr>
<td>5</td>
<td>600</td>
<td>850</td>
</tr>
<tr>
<td>6</td>
<td>850</td>
<td>1100</td>
</tr>
<tr>
<td>7</td>
<td>1100</td>
<td>1500</td>
</tr>
<tr>
<td>8</td>
<td>1500</td>
<td>2000</td>
</tr>
</tbody>
</table>

3 Mountain Wave Model

The Mountain Wave Model (MWM) presented in this paper consists of three independent parts. The first one is the identification of mountain ridges from elevation data. The use of ridges follows the approach presented in Bacmeister (1993) and Bacmeister et al. (1994), but different methods are used for the parameter determination. The second part of the model consists of the translation of the determined ridge parameters into MW launch parameters. In the third part, the so-initialized gravity waves are propagated through the background atmosphere using the ray-tracer GROGRAT (Marks and Eckermann, 1995).

3.1 Ridge Identification

The first specific task of our MWM is the estimation of MW parameters, as horizontal wavelength, amplitude, orientation and location from a fit of an idealized mountain ridge to the topography. An overview of the algorithm and examples for the intermediate steps of the ridge identification as applied to a part of South America is given in Fig. 1.

First, the elevation data is divided into overlapping slices of $10^\circ$ in latitude and every $7.5^\circ$ spanning the full globe in longitude and interpolated onto an equidistant grid in terms of physical distance. This avoids possible errors in the fitted width and length of the ridges due to projection at high latitudes. In order to identify mountain ridges of different scales, a Gaussian band-pass filter is applied as a second step. This also removes large scale plateaus, which are not of interest in this iteration of our model. In this study we are using the band-pass intervals given in Table 1.

For each band-pass filtered slice of topography, the main ridge lines (or arêtes) are identified by detecting a corresponding sign change in the gradient, which itself is calculated by a convolution with an optimized $3 \times 3$ Sobel operator (the Scharr operator, Jähne et al., 1999). To account for ridges of different orientations, this identification (and gradient calculation) is performed for 4 different directions separately (zonal, meridional and both diagonals). The result of this step is a field of line structures, representing the ridge lines (c.f. Fig. 1 d)). We determine the characteristics of these lines using a probabilistic Hough transformation, which yields the locations, lengths and orientations of lines contributing to the total structure. The
Figure 1. Flowchart of the algorithm and the intermediate steps exemplified for South America. Panel a) – Flowchart describing the main steps of the Mountain Wave Model. Input data is shown in green, internal processing steps in yellow and MWM output in blue. Panel b) - e) – example of the intermediate steps and output of the source detection algorithm. b) – input topography data, c) – bandpass-filtered topography on an equidistant grid, d) – reduction to the ridge lines and Hough transformation for a single direction and scale. Black lines represent the detected ridge lines of the bandpass filtered field, magenta are the found straight line segments from the Hough transformation. And e) – reconstruction of all fitted idealized mountain ridges.

Advantage of this approach is the sampling of arbitrary lengths and orientations by the same method. For completeness, the Hough transform as well as its probabilistic variants and their sensitivity to parameters are briefly described in the App.B.

After the line-like features in the bandpass-filtered topography have been identified, the width and height of each possible mountain ridge at the line’s location and orientation is estimated by a fit of an idealized Gaussian-shaped mountain. Note that
the location is taken from the lines found by the Hough transformation while the fit itself is performed on the bandpass-filtered topography. The cross section of the idealized Gaussian ridges is given by:

\[ f(x) = h \exp \left(-\frac{x^2}{2a^2}\right) \]  

(1)

with \( x \) being the distance perpendicular to the ridge line and \( h \) and \( a \) parameters for the mountain’s height and width, respectively. The fit minimizes the mean absolute error and performed with a 2D ridge, constructed by extending this cross section for the length of the line, which was determined by the Hough transformation.

As a result of the combined ridge finding algorithm we obtain a set of ridges with the following parameters: ridge length \( L \) and location \( X, Y \), angle with respect to the equator of the ridge \( \theta \) and the best fit width \( a \) and height \( h \).

The previous step results in a collection of mountain ridges. We assume that each of these ridges can excite a MW. In order to propagate the wave with GROGRAT we need wave vector and wave amplitude. The amplitude is taken as half the height \( h \) found from the best fitting idealized mountain, which results in the displacement amplitude \( \zeta = \frac{h}{2} \). The horizontal wave vector is chosen perpendicular to the ridge orientation and the horizontal wavelength is set to \( \lambda_{\text{hor}} = 2\pi a \), where \( a \) is the width of the best fitting idealized Gaussian ridge. This is the wavelength of maximum response for the given mountain shape (c.f. Nappo, 2012). Lastly, the vertical wavelength can be found from the dispersion relation and background data (see e.g. Fritts and Alexander (2003)), which is taken care of by the ray-tracer. An overview of all parameters estimated for each MW by the MWM is given in Tab. 2.

Using overlapping latitude slices, we sample the same topography more than once. In addition, it is not guaranteed that a single mountain ridge results in only one straight line in the Hough transformation. Thus, we expect redundancies in the ridge collection at this stage. In order to avoid double counting, we test each ridge against all other ridges by the following criteria:

- horizontal wavelength differs by less than 20%
- orientation differs by less than 22.5°
- distance parallel to the ridge is no more than 0.5L
- distance perpendicular to the ridge is no more than 0.5\( \lambda_{\text{hor}} \)

If for any ridge pair all of these criteria are fulfilled, they are assumed to describe the same underlying ridges and only the one with better fit to the bandpass-filtered topography is used. After this filtering, we end up with the final mountain ridge database.

3.2 Ray Tracer

Based on the mountain ridge database established in Sect. 3.1, this section describes the derivation of MW launch parameters for specific atmospheric conditions and details of the propagation calculation within the ray tracer GROGRAT (Marks and Eckermann, 1995). Since we are considering mountain waves, the ground based frequency of all our waves is assumed to be zero, \( \omega_{gb} = 0 \), which in turn leads to the intrinsic frequency \( \omega = -kU - lV \), where the zonal and meridional background winds,
Table 2. Parameters derived within the Mountain Wave Model – either directly from topography data or from other parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Estimation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$L$</td>
<td>Ridge length</td>
<td>from Hough transformation</td>
</tr>
<tr>
<td>$X,Y$</td>
<td>Lon, lat position</td>
<td>from Hough transformation</td>
</tr>
<tr>
<td>$\theta$</td>
<td>Math. angle of ridges w.r.t. local x-coordinate</td>
<td>from Hough transformation</td>
</tr>
<tr>
<td>$a$</td>
<td>Width of idealized ridge</td>
<td>from Gaussian fit</td>
</tr>
<tr>
<td>$h$</td>
<td>Height of idealized ridge</td>
<td>from Gaussian fit</td>
</tr>
<tr>
<td>$\lambda_{hor}$</td>
<td>Horizontal wavelength</td>
<td>$2\pi a$</td>
</tr>
<tr>
<td>$\zeta$</td>
<td>Displacement amplitude</td>
<td>$\frac{h}{2}$</td>
</tr>
</tbody>
</table>

$U$ and $V$, and the horizontal components of the wave vector, $k$ and $l$, are determined by atmospheric background data and the MWM respectively. Positive/negative $\omega$ corresponds to waves propagating against/with the wind. Since MWs are only excited propagating against the wind flow, we will assume positive $\omega$ in the following.

To account for surface friction of the low level wind and potential blocking at low speeds, a reduced, effective displacement amplitude is calculated following the discussion in Barry (2008, pp. 72-82) and assuming that the displacement amplitude of a GW excited by air forced upward is half the air parcels vertical displacement:

$$
\zeta_{eff} = \min\left(\zeta, \frac{0.32U_{par}}{N}\right),
$$

where $N$ is the Brunt-Väisälä frequency and $U_{par}$ is the horizontal wind velocity projected onto the wave vector. This means that the displacement amplitude is reduced in case the wind and stability do not allow for the full amplitude the mountain could excite. Using this effective displacement, the wind amplitude can be calculated following linear theory (e.g. Fritts and Alexander, 2003) via

$$
U_{amp} = \frac{N}{\sqrt{1 - \left(\frac{f}{\omega}\right)^2}} \zeta_{eff}.
$$

Here, $f$ is the Coriolis parameter and $\omega$ is the intrinsic frequency of the considered GW.

The ray-tracing of the excited GWs itself is performed by GROGRAT (Marks and Eckermann, 1995), which implements the ray equations derived in Lighthill (1978). Refraction and propagation are calculated using the equations

$$
\frac{dx_i}{dt} = \frac{\partial \omega}{\partial k_i} \quad \text{and} \quad \frac{dk_i}{dt} = -\frac{\partial \omega}{\partial x_i},
$$

where $x_i, k_i$ are the $i$-th component of the position- and wave-vector, respectively. The derivative $\frac{d}{dt}$ is the standard Lagrangian derivative and $\omega$ is the intrinsic frequency given by the dispersion relation

$$
\omega^2 = \frac{N^2(k^2 + l^2) + f^2(m^2 + \frac{1}{4H^2})}{k^2 + l^2 + m^2 + \frac{1}{4H^2}}.
$$
Here \((k, l, m)\) are the components of the wave-vector and \(H\) the scale height. GROGRAT takes into account background fields varying in space and time, allowing for 4D ray-tracing of GWs. The background fields (cf. Sect. 2.2) are internally interpolated to the current ray location using precalculated cubic spline coefficients. This allows for efficient calculation and avoids discontinuities in the background variables and their derivatives.

The dispersion relation in Eq. 5 is derived for small-amplitude waves in a slowly varying background flow and neglecting vertical wind. Acoustic waves are neglected for the derivation and the Boussinesq approximation is used. For the full theory, see (e.g. Nappo, 2012; Fritts and Alexander, 2003).

GROGRAT implies conservation of wave action density, \(A = \frac{E}{\omega}\), along each ray’s path except for wave saturation and for a term parameterizing radiative and turbulent damping:

\[
\frac{dA}{dt} + \nabla \cdot (Ac_g) = -\frac{2A}{\tau},
\]

where \(c_g\) is the wave’s group velocity and \(\tau\) the parameterized damping time scale.

Damping due to turbulence of the background is based on approaches presented in Hines (1960) and Pitteway and Hines (1963), depending on the inverse Prandtl number and the background diffusion coefficient (the vertical profile of which is taken from the approximation in Hocking (1991)). Radiative damping terms are taken from Zhu (1993).

In this study, we use the implementation of the saturation scheme described in Fritts and Rastogi (1985). This takes vertical dynamical (Kelvin-Helmholtz) instability, which is especially relevant for low frequency waves, and convective instability, where the waves local temperature perturbations break vertical stability, into account.

For more details on the inner workings of GROGRAT, see Marks and Eckermann (1995) and Eckermann and Marks (1997).

### 3.3 Representation of ray-tracing data

For the analysis and discussion of Sect. 4 and Sect. 5, ray-tracing data generated by the MWM is presented using two approaches: as residual temperature structures and as momentum flux distributions. The post processing steps to generate these data sets from model output are described in this section.

#### 3.3.1 Residual temperature reconstruction

For the residual temperature fields, we aim to reconstruct GWs and their temperature perturbation on a specified spatial \((x, y, z)\) grid for a selected time \(t\) for all considered rays, i.e. as a superposition of all predicted MWs. Each individual ray-trace is represented by a wave packet centered around the ray-path position of the reconstruction time. The parameters needed for this reconstruction, that is spatial location, wave-vector, phase and amplitude, are linearly interpolated to the selected time. The phase at a given point of the spatial reconstruction grid is calculated via

\[
\phi_{\text{tot}} = \phi + k_{\text{hor}}d_{\text{along}} + md_z + \frac{c_m}{2}d_z^2.
\]

Here \(k_{\text{hor}}\) and \(m\) are the horizontal and vertical wave number, \(d_{\text{along}}\) is the horizontal distance from the center of the GW in direction of the wave-vector and \(d_z\) the vertical distance of the corresponding grid point to the rays location. \(\phi\) is the current
phase at the ray-path of the wave given by the ray-tracer. The last term accounts for linear frequency modulation in the vertical
with chirp rate \( c_m = \frac{\Delta w}{\Delta z} \), which is calculated as the finite difference derivative of \( m \) for the closest time steps around target altitude \( z \).

The spatial extent of the wave packet perpendicular to the wave-vector is estimated using the length \( l \) of the ridge exciting
the MW. In this direction, the amplitude is scaled with an additional symmetric Butterworth function with \( l/2 \) as scale for a
smoother transition to zero at the edges. In direction of the horizontal wave-vector, a Gaussian-like envelope with \( \sigma = \frac{\lambda_{\text{hor}}}{2} \) is
applied. Likewise in the vertical direction a Gaussian with \( \sigma = \frac{\lambda_z}{2} \) is used. The total contribution of a single wave packet is
therefore given by

\[
T = -T_{\text{amp}} \cos(\phi_{\text{tot}}) \frac{1}{\sqrt{1 + \left(\frac{2d_{\text{perp}}}{\lambda_{\text{hor}}}\right)^2}} \exp \left( -\left(\frac{2d_{\text{along}}}{\lambda_{\text{hor}}}\right)^2 - \left(\frac{2d_{z}}{\lambda_z}\right)^2 \right),
\]

where \( T_{\text{amp}} \) is the temperature amplitude given by the ray-tracer, \( d_{\text{along}}, d_{\text{perp}} \) and \( d_z \) are the horizontal distances along and
perpendicular to the wave vector and vertical to the location given by the ray-tracer. \( \lambda_{\text{hor}} \) and \( \lambda_z \) are the interpolated horizontal
and vertical wavelengths, respectively. Note that the waves start with a cold phase directly above the mountain where the
integrated phase along the ray is \( \phi_{\text{tot}} = 0 \).

The total temperature perturbation, i.e. the total distribution of residual temperatures, is taken as the superposition of all
these individual fields for all ray-traces.

3.3.2 Momentum flux

For comparisons with satellite and other model data, GWMF of all rays has to be expressed as a superposition on a regular
spatial grid. Similar to the residual temperature, the wave parameters of the rays are interpolated to the considered time and
the momentum flux of each wave packet is distributed across the specified data grid using Eq. 8 analogously for GWMF, but
omitting the phase term and using a faster decay (by a factor of 2), since \( \text{GWMF} \sim T^2 \).

The size of a wave packet is finite compared to the 1D trajectory given by the ray-tracer and a single ray may contribute
to several grid cells of the regular target grid. On the other hand, if a grid cell is larger than the wave packet, only a fraction
of the grid cell is covered and this has to be taken into account by normalization. In order to account for this effect, we are
supersampling the GWMF of each wave on a finer grid (3 \( \times \) 3 subgrid resolution for each grid point) and averaging over the
finer grid points within each original grid box. This gives us a numerical approximation of the grid cell integrated GWMF.
Since we are interested in the GWMF density, we divide the total contribution by the grid cells area:

\[
F_{\text{tot}} = \frac{F_{\text{grid}}}{A_{\text{grid}}} = \frac{\int FdA}{A_{\text{grid}}},
\]

where \( F_{\text{tot}} \) is the wave’s contribution to the GWMF density distribution, \( F_{\text{grid}} \) is the GWMF in the corresponding grid cell, the
integration is across the (horizontal) area of the grid cell and \( A_{\text{grid}} \) is the footprint of the grid cells of the horizontal distribution.

The total GWMF density distribution is again calculated by summing up the single contributions of every ray-traced wave
packet for each time step.
4 Model Validation

In this section, we validate the MWM presented in Sect. 3 in two different ways. First, we investigate the performance of describing the underlying topography and its structures by the idealized ridges and second, we look at the resulting temperature residuals of MWM predicted mountain waves in comparison to ECMWF IFS data.

4.1 Detected structures and scales in the MWM

To investigate the capability of the MWM to adequately represent the topographic structures, we show a comparison of the underlying elevation to small (≤150 km) and large scales (>150 km) of detected ridges in Fig. 2. Although a perfect reconstruction of the underlying topography is not what we wanted to achieve with the MWM, the ridge like mountain chains should be represented by the model. Our algorithm of fitting idealized ridges of various scales and sampling in four directions implies that a single topographic feature could be captured multiple times. This is especially the case for very large scale and plateau-like features. Therefore, the elevation height of the reconstruction might be higher than the topography itself.

Higher values in absolute (superposed) terrain height does not inevitably lead to an overestimation in GW activity. For example, if there are two ridges lying on top of each other at a right angle, depending on the wind, only one of both ridges will excite a GW. Therefore, the correlation of terrain height and GW temperature amplitude is not as straightforward.

Figure 2a shows the true elevation of the terrain of South America, while panels b and c show the small and large scales detected by the MWM, respectively. The reconstructions are generated by superposing all found ridges linearly. The small scale ridges are located mainly in the main north-south oriented mountain ridge where we would expect them, but additionally there are also quite few in the east and at the tip of Tierra del Fuego with an east-west orientation. These eastern ridges can also be found in the topography itself. The general structure represented by the small scales agrees well with what we see from the topography. The large scales form a broad ridge-back along the Andes with a maximum above the highest elevation of the southern Andes around 50°S and another one to the north, where the topography shows a large scale plateau-like elevation, that is sampled along multiple directions and therefore more pronounced in the reconstruction.

In Fig. 2d, the elevation profile of the east coast of the USA including the Rocky Mountains is shown. Again, panels e and f show the detected small and large scales, respectively, as detected by the MWM. This case poses a more difficult problem for the MWM due to the topography being even more complex. While South America exhibited ridge like features reaching down to sea level from the start, the Rocky Mountains are already located on top of a very large scale elevation feature. Nevertheless, the small scale ridges represent elongated mountain chains, like for example the Sierra Nevada and Baja California, well and the structures can be matched to the topography. A more scattered structure is seen in the large scale features. Increased elevation is found above high mountain ranges, like the Rocky Mountains, Sierra Nevada and Sierra Madre Occidental. Again the topography is well represented by the identified idealized mountains.

The different scales of the detected mountain ridges and their spread are shown in Fig. 3 for South America and the Rocky Mountains. Both cases are very comparable in their distribution of ridge height, length and horizontal wavelength. The MWM detects mostly ridges with small lengths between ~75 km and 150 km. Although there are a few outliers with lengths of up
Figure 2. Comparison of the underlying topographic data (a and d) to the reconstruction from all idealized Gaussian mountain ridges identified by the MWM (b, c, e and f). Panel a, b and c show the Southern Andes region, panel d, e and f the west of North America. The reconstruction is separated into small scales (≤ 150 km, panels b and e) and large scales (>150 km, panels c and f).

to ∼500 km, a natural topography is not easily described by straight ridges due to curving or bends in the mountain structure. Thus the longer ridges are usually only detected for the largest scale, low amplitude features. Ridge heights are almost evenly distributed in a range of ∼100 m to 600 m and extending up to 800 m. The distribution is minimally shifted towards higher ridges in the Rocky Mountain case accounting for the higher elevation in this region. Considering the horizontal wavelengths of the detected ridges, they are evenly distributed between 100 km and ∼400 km with sparser distribution up to 1150 km. This can be attributed to the small scales in general being of shorter length as well and thus, they cover not as much area and multiple ridges are needed for description of the topography. The larger scale ridges on the other hand cover a large area and less of them are needed for a similar region.

In conclusion, the investigation of detected elevation features by the MWM does a good job in representing features of various scales. Although this is no indication whether we cover all relevant scales, it provides confidence in the underlying ridge detection algorithms.

4.2 Residual temperature as compared to ECMWF operational analysis data

Figure 4 compares the residual temperature due to gravity waves taken from ECMWF IFS operational analysis data on 21.09.19 at 06:00 UTC to a reconstruction estimated from MWM data. The IFS data set has been detrended in the same manner as the ERA5 data used for ray-tracing backgrounds (see Sect. 2.2), only now we are interested in the remaining residual fields representing small scale gravity wave perturbations instead of the background. The operational analysis data set is provided on a 0.1° resolution and capable of resolving mesoscale gravity waves. We estimate the smallest scale GWs resolved in the data.
Figure 3. Scales of detected ridges that contribute to the approximation of the Southern Andes region (Fig. 2 b and c) and western North America (Fig. 2 d and e). The histograms for ridge length in km (pink), ridge height in m (green) and the main horizontal wavelength in km (blue, cf. Sect. 3.1) are shown.

set to be about 80 km horizontal wavelength by the scale, where the energy spectrum deviates strongly from a slope of $k^{-3}$ (c.f. Skamarock, 2004). The reconstruction of residual temperature from the MWM is described in Sect. 3.3.1.

First, we consider the residual temperatures below the tropopause at an altitude of 8 km (Fig. 4 a and b). At this altitude, we do not yet expect strong horizontal propagation of smaller scale MWs, which typically starts to become relevant above the tropopause, thus we expect the MWM to exhibit GW patterns close to the main mountain structures. The IFS data shows high activity of all scales above the oceans as well at this height, which indicates either wave sources other than orography (e.g. convection, spontaneous adjustment, jet fronts) or completely different tropospheric processes, that our scale separation anomalously picks up. Above the Andes, we see the typical north-south oriented wave fronts and an enhanced region of small scale perturbations at around 50°S, 74°W in both data sets. The warm temperature phase follows the coastlines in both data sets. The MWM also shows trailing waves at the tip of Tierra del Fuego, which are oftentimes seen in temperature observations of this region (e.g. Preusse et al., 2002; Jiang et al., 2012; Kruse et al., 2022). There is also a large scale pattern eastward of the continent, which can be associated with a similar pattern in the IFS data. However, we cannot be sure whether the pattern in the model arises from orography. Both data sets agree in terms of the magnitude of temperature amplitudes.

At 20 km altitude (Fig. 4d and e), we expect to observe at least some oblique propagation of MWs due to stronger winds and wind gradients. Still, both data sets show the coast line following warm phase followed by trailing waves in the south. The highest amplitude residuals are found in the same location of ∼42.5°S, which is in the lee of the highest elevation topography. In terms of structure, both show a large scale pattern interfering with smaller scale GWs. The IFS, however, seemingly does a better job of resolving very small scales. Above the Falkland Islands, both show GWs that propagated towards towards the east, the GWs of the MWM are of relatively low amplitude, while the IFS shows higher amplitude and larger scale waves.
At 30 km altitude, the polar vortex turns from mainly eastward to a more north-eastward direction above the Andes. This leads to a change in the horizontal wind gradients, and in turn to GWs refracting and turning as well (Fig. 4g and h). Both data sets show gravity waves mainly facing to the south-west instead of westward, as a consequence of horizontal refraction towards the stronger winds in the south-east (Krasauskas et al., 2022). We see significant GW activity over both the Pacific Ocean as well as the Atlantic Ocean in the IFS data, which can be mostly explained by oblique propagation of MWs from the Southern Andes as indicated by the MWM data. The amplitudes of the waves predicted by the MWM agree with the IFS. The IFS model also shows large scale waves above the Pacific Ocean, which are not present in the MWM.

We analyzed the spectral characteristics of the waves around $45^\circ$-$50^\circ$S, $60^\circ$-$66^\circ$W in Fig. 4 using the S3D technique (3D wave fitting in a subdomain, previously used by e.g. Preusse et al. (2014); Geldenhuys et al. (2021); Krasauskas et al. (2022). S3D returns estimates of the 3D wavenumber $(k, l, m)$ which are then used as input for a reverse ray tracing calculation using GROGRAT. This calculation suggests that these waves originate in a region with large scale, low elevation topography leeward of the main Andes ridge ($\sim$S, $70^\circ$W in Fig. 4). Since this topography is "plateau like" it is missed by our ridge finding algorithm (not shown). It is also possible that these waves were excited by orographically-linked convection, which was observed by Worthington (e.g. 2002, 2015). These processes are currently not accounted for in the MWM. Smaller scale features in the northern part are represented correctly in the MWM both in terms of orientation and amplitude. Also, both models show small scale fluctuations at around $\sim$50$^\circ$S, 73$^\circ$W.

Figure 5c, f and i shows the wind situation at this time. We see a turning wind with altitude: while the wind is passing the Andes mostly eastward below 20 km, the wind turns to north-north-east above. Thereby GWs are refracted and the phase fronts change alignment to north-west to south-east. The turning of GWs in the MWM can be seen in Fig. 5. Although there are MWs launched with various directions between southward and westward, at around 25 km altitude, most of the MWs turn south-westward, leading to this being the dominant wave orientation. This also implies that the change in wave field characteristic is not happening due to filtering of westward facing GWs, but instead due to refracting GWs in the turning wind profile.

In comparison to the high resolution IFS simulations, the MWM does perform quite well, if we consider the simplicity of the approach. Of course, we do not expect the MWM to represent the exact structure of the IFS, since nature is more complex than this superposition of linear 2D-like MWs can account for. The aim of the MWM is to predict the horizontal propagation in a comprehensive way allowing for an straightforward investigation of propagation paths and momentum transport pattern. We have seen that the model is capable of doing so and therefore is a useful tool to estimate MW residual temperature structures. Especially the GW field characteristics, like turning and propagation of momentum, are captured quite well, and therefore the MWM can be used to investigate GW parameters of observations and propagation patterns in the following section.

5 Prediction of global GWMF distributions

In the following, we use the MWM and it’s GWMF predictions to explain some observational features of monthly mean HIRDLS satellite data via investigation of the GW parameters and wind profiles, i.e. GW filtering due to wind conditions. In particular, we try to answer the question why there is not as much GW activity seen by satellites as expected above the
Figure 4. Temperature residuals from ECMWF IFS operational analysis data for 21.09.21 06:00 UTC (left column) and the corresponding reconstruction from MWM data at the same time (middle column). Horizontal cuts are shown at altitudes of 8 km, 20 km and 30 km in the top, middle and bottom row respectively. Note the differing color scales between different altitude levels. Synoptic wind fields as predicted by the IFS are shown on the right column, where colorshading gives the total wind speed and wind barbs show the wind direction.

Himalaya and the Rocky Mountains. Afterwards, we look into predicted orographic GWMF throughout the year, which agree with previous studies of MW propagation. Here, the MWM predicts strong year-round horizontal MW propagation.

5.1 Global distributions of momentum flux

HIRDLS in general does not show strong GW activity above the Himalaya and Rocky Mountain region in January, contrary to the general observation, that mountain waves are able to propagate far into the stratosphere in the winter months. We will investigate possible reasons behind this in the following by comparing global horizontal distributions of GWMF retrieved from
Figure 5. Directional vs. altitude distribution of MWs as modeled by the MWM. The radius gives the altitude of a given MW, angle gives orientation of the wave vector. The color-shading gives the total count of rays for a given altitude-orientation bin. There is a clear trend for GWs to turn to south-west at around 25-30 km. Although the sum over all orientations for a given altitude is monotonically decreasing with altitude, there is a maximum at around 35 km with south-west oriented GWs.

HIRDLS to predictions of the MWM at altitudes of 16 km, 20 km, and 25 km. We will also compare both data sets in terms of propagation patterns found in winter above the Drake Passage and the Southern Ocean.

The distributions retrieved from HIRDLS show gaps in the lower altitudes due to the tropopause and clouds. The former removes the tropics and parts of the subtropics at 16 km and 20 km, while the latter especially removes regions of subtropical convection from the data at 16 km. The satellite observations pick up GW signatures of all sources. In winter these are mainly orography and spontaneous imbalance, in summer the main source is convection. In contrast, the MWM only shows orographic GWs, which has to be kept in mind when comparing both. In particular, this leads to possibly higher observed GWMF throughout the observations, which might not be mirrored by the MWM data. This could either suggest a different GW source, if there is no similar structure in the MWM data, or an superposition of orographic and other sources, if a feature is structurally seen in the MWM.

For a direct comparison of global GWMF distributions, we apply an observational filter that accounts for HIRDLS observation geometry (Trinh et al., 2015) to every ray of the MWM and bin the resulting GWMF values in the same way as the
HIRDLS profiles in order to derive global distributions (see Sect. 2.3). Two case studies for January and July 2006 are presented in this section.

### 5.1.1 January 2006

Figure 6 shows monthly mean total GWMF distributions for January 2006 as retrieved from HIRDLS (left column) and predicted by the MWM (right column) at altitudes of 16 km, 20 km and 25 km. The strongest pattern in the observations is found above the Himalaya and Altai Mountains (Mongolia), where we see two local maxima. The one above the Himalaya dominates at 16 km, but decreases in strength stronger than the one above the Altai mountains with increase altitudes. Therefore, at 25 km only the pattern above Mongolia remains. This pattern is also rather consistent throughout the considered altitudes. The same structural pattern is seen in the MWM, although with lower amplitudes. Features in both regions show comparable local maxima at lower levels and, at higher altitudes, the one above the Himalaya is reduced analogously to the observations.

To understand this northward shift, we can investigate the properties of GWs in the different regions within the MWM. Fig. 7 shows histograms of horizontal and vertical wavelengths at 16 km and 25 km for both regions as calculated from the MWM. We see that the horizontal wavelengths are almost the same at both altitudes and regions. Conversely, the vertical wavelengths differ strongly. The MWs of the southern region (above the Himalaya) exhibits longer vertical wavelengths, which are also strongly suppressed by the observational filter (with a cutoff at $\lambda_z = 12$ km) at 16 km altitude. Propagating upwards, they refract strongly towards short vertical wavelength due to a negative vertical gradient of zonal wind (see Fig. 9). There are two possible reasons for the GWs missing in the satellite observations at higher altitudes. For one, the vertical resolution of HIRDLS is about 1 km (Wright et al., 2009), which, in principle, allows for the detection of GWs with a vertical wavelength as low as 2-4 km. Some waves are refracting even below this vertical wavelength and can therefore not be picked up by the instrument. Another reason are strong amplitude GWs, which are often found in this region, that could completely break instead of propagating further with an amplitude reduced below the saturation limit in a strong wind vertical shear (Kaifler et al., 2015). Such a complete breakdown of GWs is currently not captured by GROGRAT simulations. The MWM could be a suitable tool to test this hypothesis in other, more specific case studies, by implementing different breaking schemes.

A very strong GW signature in the MWM is seen above the Rocky Mountains. The amplitudes are overestimated in comparison to the satellite observations, but nevertheless a similar structure is visible. In higher altitudes this signature is strongly reduced until it almost vanishes at 25 km. There is also a minor southward shift of GWMF towards California visible which is also picked up by the satellite data. The strong reduction in amplitude can be explained by similar arguments as for the Himalaya region: there is a strong negative vertical wind shear above the Rocky Mountains reducing the allowed amplitudes of GWs strongly. Compared to the Himalaya region, in the MWM the MWs start out with much stronger amplitudes (about a factor of two, not shown), making them more likely to encounter saturation or complete breakdown (there is plenty of evidence of strong MWs and their breaking in this region, e.g. Guarino et al. (2018)). The latter process might be a reason for the overestimation at 16 km (the negative wind shear starts already at roughly 10 km, where waves could already break). Therefore the strong signature in the MWM data could be another hint at this or another process missing in our current understanding of GW physics.
Figure 6. Monthly mean of global GWMF distribution for January 2006. Left column shows HIRDLS satellite data (see Sect. 2.3), right column shows distributions produced by the MWM. The different rows present data for 16 km, 20 km and 25 km altitude from top to bottom respectively.

Another strong feature predicted by the MWM are local maxima in the Southern Hemisphere above New Zealand and the southern Andes, which are matched by the observations. These maxima strongly decrease in higher altitudes and vanish completely at 25 km as expected, since MWs are filtered by the wind reversal at around 20 km in the summer hemisphere. The prediction shows strong eastward propagation above New Zealand, while the satellite data seems to show signs of strong oblique propagation eastward of both sources. However, their persistence at 25 km altitude in the observations is not consistent with them being MWs.

In the northern Atlantic region, the MWM predicts GW sources in Newfoundland, southern Greenland, Iceland and Scandinavia. Eastward propagation is seen especially above Iceland, where the GWMF reaches to the patch above Scandinavia. The observation shows a more complex picture with the same sources: the single features seem to merge into each other creating a continuous band of GWMF above the Atlantic, which persists into higher altitudes. At 25 km there is increased GWMF east of Newfoundland (compared to lower levels) hinting at GWs from sources like the jet or storms above the Atlantic ocean. Most features vanish at this altitude and only the ones close to orography and eastward of Iceland remain. In comparison to the MWM, observed features at 16 km can therefore be attributed to orographic origin, while the ones observed at 25 km should have other sources (see e.g. Geldenhuys et al., 2021).

To further investigate the reasons for the differences between the MWM and the satellite observations, we use blocking diagrams similar to the ones introduced in Taylor et al. (1993) as well as vertical profiles of horizontal wind. The regions of interest are shown in Fig. 8. We investigate differences in propagation conditions above the Himalaya compared to Mongolia,
above the Rocky Mountains and above southern Africa, where a strong GWMF pattern arises in the HIRDLS data at 25 km altitude (see Fig. 6). For illustration of the general wind conditions, Fig. 8 shows the monthly mean zonal wind at 20 km.

In the blocking diagrams shown in Fig. 9a – d, the criterion of waves blocking whenever the intrinsic frequency of a GW goes to zero, $\omega_{\text{intr}} \rightarrow 0$, is used, with:

$$\omega_{\text{intr}} = \omega_{gb} - k_{\text{hor}} \cdot U$$

$$\Leftrightarrow \omega_{\text{intr}} = \omega_{gb} \left( 1 - \frac{U_{\text{par}}}{v_{\text{ph}}} \right).$$

Here, $k_{\text{hor}}$ and $U$ are the horizontal wave- and wind-vector, $U_{\text{par}}$ the wind speed projected onto the horizontal wave-vector and $v_{\text{ph}}$, the horizontal phase speed of the GW.

The curve of $\omega_{\text{intr}} = 0$ is a circle in phase speed diagrams with center at $(\frac{U}{2}, \frac{V}{2})$ and radius $R = \frac{1}{2} \sqrt{U^2 + V^2}$ for zonal and meridional background winds $U$ and $V$. These contours have been superposed from the surface to 25 km altitude and divided by the number of contours to give a measure of how strong or wide spread across heights the blocking is. Fig. 9a – d show therefore the percentage of altitude at which blocking occurs for a given (ground based) phase speed in color shading. Note, however, that these diagrams are only an indication of blocking for MWs with $\omega_{gb} \approx 0$, which are blocked wherever the horizontal wind projected onto the horizontal wave-vector becomes zero. As an additional metric, the monthly mean vertical profiles of horizontal wind for the four regions are shown in Fig. 9e and f.

Comparing the blocking patterns for the Himalaya and Mongolian plateau in monthly mean winds in Fig. 9a and c, we see that a wider range of phase speeds is restricted by blocking in the southern region, above the Himalaya. This potentially leads to a stronger suppression of GWs of other origin than orography, which might be part of why the GWMF in HIRDLS declines so strongly in Fig. 6. If we look at the vertical profiles of monthly mean winds, which are shown in Fig. 9e, neither of the
regions show a wind reversal and thus MWs should in general propagate similarly well in both regions. However, the zonal wind of the southern region exhibits a strong vertical gradient. The wind speed peaks at around 12-13 km with a pronounced maximum of about \(45 \text{ ms}^{-1}\), that is roughly 5 km below the level of maximum wind speed of the northern, Mongolian, region. This high wind speed allows for MWs of higher amplitudes, which afterwards encounter a strong negative wind shear leading to refraction towards small vertical wavelengths and wave breaking due to reaching the saturation limit. Kaifler et al. (2015) suggests that high amplitude GWs reaching saturation might break completely instead of propagating further with a reduced amplitude. This total breakdown of saturating waves is not represented in GROGRAT simulations, but could be a reason why HIRDLS sees less activity above the Himalaya (Fig. 6). Above the Altai mountains, the propagation conditions are more favorable due to a more consistently strong wind, that is slowing down only above 25 km. Therefore the observations and the model do not see as strong a reduction in activity as above the Himalaya region.

Next, we want to investigate the Rocky Mountains. The blocking diagram for this region is shown in Fig. 9b and corresponding wind profiles are given in Fig. 9 f (blue lines). The blocking diagram exhibits high values at and around the origin indicating low wind speeds at various heights. The dumbbell-like shape with structures on either side of the origin are a consequence of a wind reversal. This is confirmed in the wind profiles, where we see the reversal of horizontal winds at about 22 km. This reversal prevents any MW activity to propagate further upward. A possible reason for the wind reversal despite being in the winter hemisphere is the location of the polar jet. In Fig. 8 we see that the polar jet in the monthly mean is not located above the Rocky Mountains but much further north (about \(75^\circ\)N), which strongly affects the propagation criteria for MWs. In terms of shape of the wind profile, the situation is similar to the Himalaya region with a peak wind speed of about \(257, \text{ms}^{-1}\) at \(\sim\)11 km altitude followed by a strong negative wind shear. In addition, the low level wind speeds are around \(6 \text{ ms}^{-1}\), which is four times as strong as in the Himalaya region. The MWs of this region therefore have the potential to launch with much higher amplitudes, making them more likely to reach saturation and complete breaking due to high amplitudes as described by Kaifler et al. (2015). Since this is not represented in GROGRAT, the MWM may predict much stronger GWMF in this region, than there actually is in reality.
Lastly, we want to briefly consider southern Africa to exclude MWs as a source for the pattern seen in HIRDLS at 25 km. The corresponding blocking and wind profiles are shown in Fig. 9d and f (yellow lines). The blocking diagram has a pronounced dumbbell shape from a wind reversal, which is confirmed by the wind profiles. This makes propagation of MWs impossible. Conversely, only small parts of the phase speed spectrum is blocked. These are ideal conditions for GWs of sources that generate a wide range of phase speeds like convection (e.g. Salby and Garcia, 1987; Alexander and Dunkerton, 1999; Preusse et al., 2001; Choi and Chun, 2011; Trinh et al., 2016). Therefore we can conclude that the observed pattern appears not due to orography but other sources (most likely convection).

To summarize, the predictions of the MWM and the calculated wave parameters can explain the shift of focus of GW activity from the Himalaya to the Altai Mountains and therefore solve the question, why there is not as much GW activity as would be expected from the topography itself. The MWM shows, that parts of the GW spectrum refract to very short vertical wavelengths, which makes them hard to be detect by the satellite. The feature above the Himalaya is more pronounced without the instrument specific observational filter. On the other hand, the MWM is showing a strong signature above the Rocky Mountains, as would be expected from the topography. Since this is not present to such a degree in the satellite data, it might be a hint at a GW process, that is not captured as of now (e.g. total breakdown of GWs reaching saturation). In total the MWM has proven to be a useful tool to investigate the orographic part of GWMF observations.

5.1.2 July 2006

In the following, we are considering predictions of GWMF for July 2006 and use calculated MW parameters (mainly the wave vector) to explain features found in HIRDLS observations. Fig. 10 shows global horizontal distributions of GWMF as retrieved from HIRDLS (left column) and predicted from the MWM after application of the observational filter (right column) at altitudes of 16, 20 and 25 km. The main feature in both data sets is the maximum above the southern Andes. At 16 km HIRDLS shows a strong maximum at around $\sim 52^\circ$S accompanied by a weaker separate maximum directly north at around $\sim 42^\circ$S. In higher altitudes, the southern maximum vanishes, while the northern persists. In total, this leads to a northward shift of the global maximum. The HIRDLS data set also shows eastward propagation of GWMF from these maxima up to about $30^\circ$W over the Atlantic. This is a hint at orographic GWs from the southern Andes contributing in this region. Although in a textbook case, we would expect the extent of the lateral propagation to increase by height, the data set shows a reduced trail at 20 km altitude and an increase above. The MWM in comparison shows a very similar pattern with a larger maximum above the southern Andes. Although the MWM does not show a clear maximum in the southern part, the pattern is rather spread out and shifted northward at higher altitudes as in the observations. In terms of horizontal GW propagation, the data shows stronger extent of the propagation the higher one looks at. At 25 km, the GWMF is transported up to $30^\circ$W as in the satellite observation.

Using the MWM allows for an in depth look of what is causing the northward shift seen in observations by investigating individual GW parameters in these regions. Fig. 11 shows histograms of vertical and horizontal wavelengths of all GWs between $37^\circ$-47$^\circ$S and $47^\circ$-57$^\circ$S and longitudes between $40^\circ$-90$^\circ$W at 16 km and 25 km. This indicates at a strong increase in vertical wavelength in the southern part, while the northern part remains almost unchanged. Horizontal wavelengths remain
Figure 9. Blocking diagrams as introduced in Taylor et al. (1993) for the four regions shown in Fig. 8: a) Himalaya, b) Brazil, c) Mongolian plateau and d) southern Africa. Percentage of altitude levels at which blocking occurs of a gravity wave with the corresponding phase speed given on the axes is shown in color shading. These percentages are calculated as a mean over all vertical wind profiles of the regions from surface to 25 km altitude. Panels e) and f) show the monthly mean vertical profiles of zonal (solid) and meridional (dashed) wind for the Himalaya and Mongolian Plateau regions and Brazil and southern Africa region, respectively (colors correspond to the regions in Fig. 8).

mostly unchanged in both regions. Therefore in the southern part, the GWs refract towards vertical wavelengths larger than the cutoff wavelength of 12 km used in the generation of the HIRDLS data set and are thus filtered out at higher altitudes. This
finding is confirmed by the unfiltered observation data (without the cutoff at \( \lambda_z = 12 \text{ km} \)), which shows a broad maximum at all altitudes (not shown). In addition, the MWM shows more GW activity in the south without the observational filter, which could mean that HIRDLS picks up more of the horizontal spectrum than we assume here (c.f. .13).

Another major predicted feature is strong GW activity around the Antarctic Peninsula and eastward trailing GWMF, especially at higher altitudes. These predictions agree with observations in terms of extent of the horizontal propagation. The HIRDLS data, however, shows another peculiar pattern: GWMF is increasing again above 20 km altitude. Since the dataset is limited to about 63\(^\circ\)S, it is not clear where this is coming from. One source could be orographic waves propagating from further south. This is not seen in the MWM though. Even without the observational filter, there is some northward propagation above the Antarctic Peninsula, but it is far to little to compensate the reduction in orographic GWMF with altitude. Therefore, we can attribute this feature in the observations to MWs due to katabatic flow (Watanabe et al., 2006) or other, non-orographic processes like spontaneous imbalances of the polar jet or . The eastward propagation seen at 25 km in both data sets is in agreement previous studies of e.g. Sato et al. (2012).

Another predicted pattern found in both data sets is enhanced GW activity above the Southern Alps and the Great Dividing Range/Tasmania (a study of this as a MW source has been done by Eckermann and Wu (2012)). The location and strength of local maxima fits nicely between the two. In the observations, these features are of comparable strength as the background and thus difficult to disentangle. At higher altitudes, MW from Australia/Tasmania are reduced significantly and rather localized. MWs excited by the Southern Alps on the other hand show strong eastward oblique propagation already at 16 km altitude. In the observations, it is not clear whether the long south-eastward trail from New Zealand is caused by MW propagation or a different GW source, however looking into unfiltered MWM data shows a strong north-eastward propagation of MWs from eastern Antarctica, which can partly explain enhanced fluxes.

Another predicted feature are MWs of similar strength to the ones above New Zealand above South Africa with propagation towards the south-east. This region is usually not regarded as a hot-spot for GWs, however, we will show that this feature is fairly consistent throughout SH winter in Sect. 5.3. In the satellite data, this feature is obscured by the belt of high GWMF in the Southern Ocean, but could be interpreted as the bend towards the south-east of Africa, that is seen in this belt.

The MWM shows few features in the Northern Hemisphere, although only at 16 km. Above this altitude, MWs encounter a critical level and are filtered out due to the summer wind reversal. Weak GW activity above the northern Rocky mountains, Greenland and the Japanese Sea can be assigned to structurally agreeing counterparts in HIRDLS. Note however that the baseline of HIRDLS is much higher than the predicted GWMF of the MWM and these features might as well not be visible in the observations at all.

### 5.2 Zonal mean momentum flux distributions

Figure 12 shows monthly and zonal mean of total GWMF as observed by HIRDLS (a, b) and predicted by the MWM (c – f) for January (left column) and July (right column) 2006. Panels a – d show the height range limited by the HIRDLS data set (14 - 25 km), while panels e and f extend further from 10 - 30 km. The MWM considers only GWMF of orographic origin, while it was shown in Sect. 5.1 that HIRDLS measured high background fluxes all around the globe including all possible GW
Figure 10. Same as Fig. 6 but for July 2006.

Figure 11. Distribution of vertical (left) and horizontal (right) wavelengths as found by the MWM at altitudes of 16 km (top) and 25 km (bottom). The northern region corresponds to 37°-47°S and the southern region to 47°-57°S, both between 40°-90°W. The vertical red line in panels a) and c) marks the cutoff wavelength of $\lambda_z = 12$ km for the present HIRDLS data.

sources (see Fig. 6 and Fig. 10). In the zonal mean, this leads to rather high values compared to the MWM prediction, which shows gaps above regions, where no MWs are present (e.g. oceans, tropics at high altitude). This leads to the MWM zonal means being of reduced values, which is why the comparison between the data sets is focused on the structural features. The MWM gives information about which patterns in the observations are caused by MWs and which are of other origin. Mainly in the tropics, the zonal cross sections show gaps in the HIRLDS data due to clouds and the tropopause seen in Sect. 5.1. Black contour lines show the monthly mean zonal wind.
For January 2006, shown in Fig. 12a and c, the same features as in the discussion of Sect. 5.1.1 can be recognized in the zonal mean. A local maximum at around 35°N at low altitudes, which moves northward to about 50°N at 20 km corresponds mainly to the strong GWMF above the Himalaya and Altai mountains and the associated northward shift with increased altitude. The Southern Hemisphere shows most dominantly the southern Andes at around 45°S. Another low altitude maximum at 30°S in the observations is probably not robust, as Fig. 6 shows, that it stems from a very small region with few data points west of Australia. The predicted GWMF is mostly restricted to below the wind reversal, although there is a part of the Andean MWs, that reach up to 25 km at around 40°S. This can be attributed to their horizontal propagation and refraction, to circumvent the wind blocking (also complete wind blocking would require zero meridional wind as well).

Corresponding data for July 2006 is shown in Fig. 12b and d. As in Sect. 5.1.2, the dominant feature is the GWMF above the southern Andes and the Southern Ocean around 40°-50°S. The prediction shows enhanced GWMF around the neck region at roughly 16 km, 30°S. A quick check against the MWM without the observational filter (12f) shows that this is most likely not due to propagation, but due to shift of MW parameters towards better observable values. Below about 20 km, there is also southward propagation towards 60°S. Observations do not show this explicitly due to the strong band of GWMF obscuring individual features in the zonal mean. The predicted zonal mean does also show no hint of northward propagation of MWs from Antarctica, that could lead to the enhanced fluxes in the southernmost observations at highest altitudes. As mentioned in Sect. 5.1.2, this feature is therefore most likely of non-orographic origin and generated by sources higher up in the atmosphere like jet fronts and spontaneous adjustment. In the Northern Hemisphere, the strongest MW activity is seen above 65°N and can be attributed to Greenland, followed by MWs excited above the Himalaya around 40°N. Both features are confirmed by observations. The MW activity is well confined by the wind reversal in the summer hemisphere.

In order to discuss the effect of the observational filter, raw MWM data without any filtering applied is shown in Fig. 12e and f for January and July respectively. In general, a reduction of absolute values by a factor of roughly 3 - 10 due to the observational can be seen, which depends on the wavelengths and the wave’s orientation to the satellite track. Nevertheless, there are also structural changes. For January, the maximum above the Himalaya gets reduced and therefore a net northward shift can be seen. Without the observational filter, there is also an additional feature in the tropics extending quite high. At high latitudes, the Maximum above the Antarctic Peninsula and another smaller one in the Arctic are strongly suppressed by the observational filter. This stronger effect at high latitudes is due to the hard cutoff in vertical wavelength at 12 km in the observational filter. In high latitudes, waves of longer vertical wavelength are more dominant since the intrinsic frequency is usually higher due to being confined to \( N^2 > \omega^2 > f^2 \). The vertical wave number in mid frequency approximation given as \( |m| = N \frac{k_{hor}}{\omega} \), where \( N \) is the Brunt-Väisälä frequency and \( k_{hor} \) the horizontal wave number. Thus at higher latitudes, where \( |f| \) is larger, GWs have in general larger minimal frequencies leading to higher minimal vertical wavelengths.

July shows a very similar picture: the Antarctic Peninsula is strongly suppressed and the maximum between 30°-50°S is reduced in width. Also the gap at 60°S is increasingly closing at higher altitudes, while the opposite can be seen after application of the observational filter. This gap is closing even more at higher altitudes than shown. Most of the GWMF propagating into this region originates at the Antarctic Peninsula, but also smaller islands like the Falkland Islands, South Georgia and the Kerguelen Islands as well as the southern Andes contribute.
Figure 12. Monthly and zonal mean GWMF for January (left column) and July (right column) 2006. Panels a) and b) show HIRDLS data, panels c) and d) MWM data after application of the observational filter and panels e) and f) MWM data without any filtering. Contour lines show the monthly and zonal mean zonal wind for the corresponding month. Note that panels e and f show a wider altitude range than a – d.

In both cases, the MWM predicts a comparable strong feature just north of the equator, reaching up as high as 20 km in January. Since the observations are limited due to clouds, this is only seen in the model data, but since this feature almost completely vanishes after application of the observational filter, HIRDLS would probably not have observed this. The origin of the MWs is around Thailand/Malaysia and the Philippines. Another feature, that is worth further investigation are MWs crossing the wind blocking (e.g. around 40°S in January). The MWM can be used in an in-depth study, to find the pathways and conditions of these MWs. This is beyond the scope of this model overview study though.

5.3 Time evolution of GWMF distributions

Now we apply the MWM to quantify oblique propagation in more detail and get first insights of GWMF transport patterns. As presented, model allows to investigate the evolution of MW activity throughout the year and related seasonal patterns. Fig. 13 shows horizontal monthly mean GWMF distributions for January to December 2006 at 25 km altitude. Note however that the observational filter used in Sect. 5.1 was not applied to this data, a comparison to Fig. 6 and Fig. 10 can therefore give an indication of the effect of the observational filter.

At 25 km we expect already strong oblique propagation as this is usually happening in a relatively small layer above the tropopause. The summer wind reversal is also below this level. Most MWs in the summer hemisphere are therefore filtered out.
The commonly expected features are evident in the time series: during summer time, MW activity is strongly reduced due to wind filtering at the stratospheric wind reversal. This corresponds to April to September in the northern and December to February in the Southern Hemisphere. In the Northern Hemisphere, however, we see an interesting feature: the MW activity above Asia is suppressed for a longer time period (April to September) than above the Rocky mountains and in high latitudes (May to approximately August). The reasons behind this could be the position of the wind reversal and might only be a feature special to 2006, which is considered here.

The data shows a band of high GWMF above the northern Atlantic connecting New England to Greenland, Iceland and Europe, which is only interrupted from May to September. The same pattern is visible in HIRDLS observations for January (Fig. 6 a) but was reduced after applying the observational filter to the MWM data. The total maximum of GWMF throughout the year is located in the Southern Andes and Antarctic Peninsula region, as the current understanding would suggest. There is strong lateral propagation from this region eastward up to the zero meridian, which is in agreement with the findings in Sato et al. (2012). This zonal propagation was not seen to the same extent in the previous filtered MWM data at 25 km altitude (Fig. 10) due to the rather small strength and specific wave characteristics. The eastward transport of GWMF is an all year round phenomenon with similar shape in all months though absolute values vary with season and are strongest in austral winter. An interesting finding is additional propagation westward of the Andes, which occurs throughout the year with maximum extent in October and November. This seemingly leeward propagation has been observed by the 2019 SouthTRAC campaign (Rapp et al., 2021) and extensively investigated by Krasauskas et al. (2022).

A feature not seen in the filtered data for austral winter (Fig. 10) is a stretched band of GWMF starting at South Africa reaching to the Kerguelen Islands and beyond towards Antarctica. This season in general shows the contribution of smaller oceanic islands, like the Kerguelen, South Georgia and the Falkland Islands, for generating GW activity at around 60°S and thus slowing down the southern polar vortex in GCMs. These enhanced fluxes above the Kerguelen Islands are also seen in the HIRDLS observations for (austral) autumn months (not shown), while it is overshadowed by the large scale globe-wrapping GWMF band for July shown in Fig. 10. The distributions agree to the findings of Perrett et al. (2021), which state that GWs above smaller oceanic islands tend to propagate shorter horizontal distances compared to the ones above the Southern Andes and Antarctic Peninsula. The strong MW activity and propagation seen above South Africa is a new finding, that has not been considered before. Due to it’s persistence for about 4 months, it has a significant impact on the atmosphere.

In general, the MWM predicts strong horizontal propagation all around the globe, especially in the Southern Andes - Drake Passage - Antarctic Peninsula region and New Zealand. But also other parts, that are not as prominently known for it, like the smaller oceanic islands, northern Atlantic and South Africa, should not be neglected. Another point to be made is the similarity and consistency of propagation patterns. For example, if there is strong activity in the Southern Hemisphere, it will be advected towards the west. This might be a general pattern, that approximates a large part of MW propagation throughout the year. Such a simple, flow independent transport pattern could be used for improving MW parameterizations in GCMs.
6 Conclusions

In this study, we present a straightforward approach and model for the localization and quantification of orographic gravity wave sources. Using a similar approach as Bacmeister (1993), we use a fit of idealized Gaussian mountain ridges to topographic elevation data for an approximation of the main orographic features able to excite mountain waves. These Gaussian ridges allow for an estimation of the main MW parameters of horizontal wavelength, orientation and amplitude. We show that our model is able to represent the general ridge structures found in topographic data. Plateaus and largest scale features (which could also lead to MWs by katabatic flow) are not considered in the current state.

Using the ray-tracer GROGRAT (Marks and Eckermann, 1995), we quantify vertical and oblique propagation as well as refraction of excited MWs within the model in time-dependent background atmospheres. Our results show that the MWM is capable of reproducing residual temperature structures comparable to the high resolution ECMWF IFS operational analysis. Though the agreement is only qualitative, the ones with expected orographic origin are well reproduced regarding orientation, scale and amplitude. In particular, smaller scale features agree in location and amplitude between both data sets. Comparisons of

Figure 13. Horizontal monthly mean GWMF distribution for each month of 2006 at 25 km altitude. Shown is MWM data without the observational filter applied.
The degree of horizontal propagation is compatible with results of previous studies of Sato et al. (2012). The agreement is therefore adequate for our subsequent MWM based analyses.

Investigations of the evolution of MW parameters show that some of the waves refract to very long (Southern Andes) or very short (Himalaya) vertical wavelengths and thus move out of the visibility range of the HIRDLS data set presented here. Refraction leads to a northward shift of the maximum above the Southern Andes and a reduced signal above the Himalaya. In addition, a study of blocking and wind filtering shows that GWs in general find better vertical propagation conditions above Mongolia compared to Himalaya. This is associated with a stronger suppression of GW (and MW) activity above the Himalaya.

The wind and propagation conditions above the Rocky mountains show a complex situation for January 2006. The low level winds are very strong (a factor of 4 stronger than in the Himalaya region) and allow therefore for excitation of strong amplitude MWs. In the lower stratosphere the excited waves encounter a strong positive vertical gradient in horizontal wind, allowing the waves amplitudes to grow before they reach a strong negative wind shear level above leading up to a wind reversal. In this strong negative shear, high amplitude MWs reach saturation in our model and propagate further at saturation amplitude. However, it would also be possible that they break completely and instead deposit all momentum locally (Kaifler et al., 2015), a process that is currently not captured by the MWM. If this process is physical, the consequence would be that we find an overestimation of GW activity once saturation of these strong MWs sets in. One implication would be that the resulting GW drag will be predicted in the wrong altitude by the ray-tracer, although it is not certain how much energy of such a breaking wave would be redistributed to other processes like e.g. secondary wave generation.

The interpretation of comparisons of monthly and zonal mean GWMF in the height range of 14-25 km to corresponding HIRDLS data proved to be difficult due to systematic differences between the data sets, like the overlapping signals of non-orographic gravity waves in the observations. Nevertheless, we obtain good agreement in terms of wave structures for regions of high topographic wave activity. Another finding is that it could be worthwhile to implement katabatic MWs in order to obtain increased fluxes northward of Antarctica and southward of Greenland. MWM predictions for July 2006 without observational filtering show, that the gap at 60°S is closing mainly by MWs propagating from the Antarctic Peninsula towards the Drake Passage, but also from contributions of smaller islands (Falkland Islands, South Georgia and Kerguelen Islands).

Global monthly mean GWMF distributions throughout the year show that the oblique propagation of MWs is not only a seasonal phenomenon but important during the whole year. Especially MWs excited in the Southern Andes and Antarctic Peninsula propagate strongly for most of the year, both east- and westward. But also in other regions, e.g. in the Northern Atlantic, horizontal relocation of momentum by MWs is important for a large part of the year (about seven months in 2006). This part of our study also underlines the importance of smaller islands around 60°S for the GWMF budget and the need for a better representation of GWs in GCMs.
One of the main goals of this study is the quantification of oblique mountain wave propagation. The results show, that the presented approach is well suited to shed light on the behavior of MWs and their appearance in observation data (especially since the MWM provide access to the parameters of each launched GW). The model might also help in disentangling the influence of primary mountain waves (which are the only waves our model is considering) and waves of other sources (also secondary waves). Another sign for reasonable representation of oblique propagation is the gap at 60°S closing at higher altitudes, which is partially seen in the zonal mean GWMF comparison of Sect. 5.1 and Sect. 5.2.

The results of this study suggest that MWs generated by katabatic flow as well as isotropic mountains (e.g. smaller islands) are worth investigation and inclusion in the MWM for it to capture all possible orographic sources. However, the MWM’s performance is good and comparisons of the model’s residual temperature to observation campaigns are worthwhile as another tool for explaining the measured GWs path and origin. Thereby the model can be used to separate MWs from non-orographic GWs in observations. Our results support that horizontal propagation of MWs is a strong and global effect that stays important throughout the whole year and is currently not considered in lower resolution GCMs.

Since there is a good agreement of the wave field characteristics of the MWM to observations and more sophisticated models, it can be used as a predictor for the momentum transport of MWs. In particular, we have shown that there is a more or less general pattern of GWMF redistribution throughout the year, which can be used as a first order approximation of the horizontal momentum relocation in GCMs. Due to the implementation using a ray-tracer, all necessary information, such as location, momentum and scales, are covered by the MWM. Implementation of such a pattern into a GCM could improve the predictability, especially in the polar vortex region.

Data availability. HIRDLS Level-2 data are available via the NASA Goddard Earth Sciences Data and Information Services Center (GES DISC) at https://acdisc.gesdisc.eosdis.nasa.gov/data/AuraHIRDLSLevel2/ (NASA GES DISC, 2022)

Code and data availability. Access to the code and data is possible upon request to the authors.

Appendix A: Mountain Wave Model Details

A1 Bandpass filter

The bandpass filter in use in the MWM is based on convolutions with a Gaussian function kernel. Given the small and large scale bounds \( \lambda_{\text{small}} \) and \( \lambda_{\text{large}} \), the corresponding \( \sigma \) of the kernels is calculated depending on the grid spacing of the topography \( (d_x) \) via:

\[
\sigma_i = \frac{\lambda_i}{6d_x}.
\]
In this way both scales are converted to a pixel scale, which is then used for the smoothing. For the final bandpass filtered topography $H_{\text{bandpass}}$, we subtract the large scales from the small scale smoothed field:

$$H_{\text{bandpass}} = H * G(\sigma_{\text{small}}) - H * G(\sigma_{\text{large}}).$$  \hspace{1cm} (A2)

**A2 Mountain wave fit**

The least squares fit is performed on a rectangular cutout of the bandpass filtered topography, $H_{\text{clip}}$. The length of this cutout is given by the line length of the Hough transformation, the width is set to $\lambda_{\text{large}}$, the upper cutoff of the bandpass filter. The lowest point of the cutout is set to zero via

$$\tilde{H}_{\text{clip}} = H_{\text{clip}} - \min(H_{\text{clip}}).$$ \hspace{1cm} (A3)

Afterwards, we fit the idealized ridge, $R$, to the cutout using the Frobenius (or Euclidean) norm and minimizing the cost function

$$f_{\text{cost}} = \frac{1}{\text{size}(H_{\text{clip}})} \text{norm} \left( \frac{R - \tilde{H}_{\text{clip}}}{\max(H_{\text{clip}})} \right).$$ \hspace{1cm} (A4)

We divide by the maximum of the cutout topography to get the relative error and by the size of it (meaning the number of grid points) to prevent overvaluing small scale ridges. This cost function represents thus the mean relative squared deviation of the idealized ridge to the bandpass filtered topography.

**Appendix B: The (Probabilistic) Hough Transformation**

The here used Hough Transformation and it’s probabilistic variant are described in the following. In addition, the sensitivity on parameters and their choice for this study is given.

**B1 Algorithm and Parameter Choice**

The Hough transformation can be interpreted as a discretized version of the Radon transformation, which is widely used in tomography. It aims at representing a 2D image (or 2D data in general) by a superposition of (a relatively low number of) straight lines. It explicitly relies on the fact that all lines passing through a point $X = (x, y)$ can be described by $R = x \cos(\theta) + y \sin(\theta)$ for $\theta \in [0, \pi)$, with the line’s distance to the origin $R$ and inclination $\theta$. The first step of the Hough transformation is the initialization of an empty accumulator $Hf(R, \theta)$ with discrete $R \in [-r_{\text{max}}, r_{\text{max}}]$ and $\theta \in [0, \pi)$. The actual size of this matrix depends on the prescribed resolution in $R$ and $\theta$ and on $r_{\text{max}} = L_x + L_y$, with $L_x$ and $L_y$ being the width and height of the input image. Now, for each non-zero pixel of the image $X_j = (x_j, y_j)$ and each orientation $\theta_i$ a value for $R_{j,i}$ can be calculated. Afterwards, the corresponding point in the Hough space accumulator $Hf(R_{j,i}, \theta_i)$ is incremented by one (that pixel in the input image basically gives a ‘vote’ for all straight lines passing through it). After processing all non-zero pixels of the image, this will result in local maxima in the accumulator $Hf(r, \theta)$ at exactly those points that correspond to main line-like structures in the image.
The probabilistic Hough transformation improves upon this algorithm by not explicitly processing every possible line through every non-zero point. Once a single line in the Hough space accumulator reaches a given threshold, the corresponding line is scanned along its slope for further connected points until a gap of point of length $l_{\text{gap}}$ or the image’s border is encountered in either direction. This gives both endpoints of the corresponding line segment and thus fully localizes it within the image. Now, if the traversed line is longer than a given minimal length $l_{\text{min}}$, it is accepted and the points making up the line are no longer considered for further iterations. The initial set of points is processed in a probabilistic order, hence the name.

B2 Sensitivity of the probabilistic Hough Transform

As stated in the main text, the probabilistic Hough transforms capability to detect mountain ridges is highly affected by the minimal line length, $l_{\text{min}}$, and the maximal gap along each line, $l_{\text{gap}}$. This is especially true for case of this work, since natural ridges typically do not form perfect straight lines, but arc or squiggle. Thus, this section investigates the sensitivity of the Hough transform on these parameters in the case of the Southern Andes.

To this end, we tested different combinations of $l_{\text{min}}$ and $l_{\text{gap}}$ and displayed the results in Fig.B1. For small $l_{\text{min}}$ and $l_{\text{gap}}$ (Fig. B1 b)), the structures are well covered by detected lines. However, these lines are in general very short and thus would not correspond to ideal 2D-like mountain ridges. For higher $l_{\text{min}}$ in combination with a small $l_{\text{gap}}$ (B1 c)), only the straightest structures are detected, which gives good candidates for ideal long stretched mountain ridges, but neglects possible smaller scale ones. A too small $l_{\text{gap}}$ combined with a high $l_{\text{min}}$ (Fig. B1 e)) is highly restrictive, and does no good job in describing the general patterns. A higher $l_{\text{gap}}$ (Fig. B1 c), f)) on the other hand leads to artifacts crossing large gaps between true structures.

Since there is a need to also detect natural curved and non-perfect straight mountain ridges, one needs to balance both parameters. There is certainly a trade off, if one wants to detect mountain ridges of arbitrary length with a single set of $l_{\text{min}}$ and $l_{\text{gap}}$.

We found that $l_{\text{min}} = 60 \text{ km}$ and $l_{\text{gap}} = 30 \text{ km}$ results in a reasonable trade-off for both parameters. With a minimal line length of 60 km most of the features are detectable within the scale ranges that we are interested in. This choice of parameters is used for all results in the present study.

Appendix C: Alternative representation of ridges as used in previous studies

The here presented mountain wave model is inspired by the study in Bacmeister (1993) and Bacmeister et al. (1994). To compare it to these previous mountain wave modeling studies, we have recreated similar plots to their representation of mountain ridges. For this we approximate each mountain ridge with a rectangular box, where the rectangles width is given by half the estimated wavelength and its length is taken as the length of the underlying ridge. As in the mentioned previous studies, we draw these rectangles atop of each other in order of ascending ridge height. This reconstruction, which is better comparable to the studies in Bacmeister et al. (1994); Eckermann et al. (2000); Kruse et al. (2021), is shown in Fig. C1 for North America including the dominant Rocky Mountain region and in Fig. C2 for the Southern Andes area, which was considered oftentimes
Figure B1. Line structure of the Southern Andes region after applying a bandpass of scales 100-250 km. Panel a) shows the underlying structure that is the input to the probabilistic Hough transform. Panel b) – f) show the detected lines (magenta) for different sets of \((l_{\text{min}}, l_{\text{gap}})\): b) – (50 km, 20 km), c) – (60 km, 20 km), d) – (60 km, 40 km), e) – (100 km, 30 km) and f) – (100 km, 50 km).

in the present study. Note that the scale of the plots shows only the height of the singular drawn ridge and not the accumulated orographic height of all ridges combined as in Fig. 2.

Author contributions. SR and PP conceptualized the study, ME performed HIRDLS processing and analysis, SR conceptualized and developed the MWM and performed the simulations. PP and MR supervised the study and JU acquired the funding. All authors provided scientific input and reviewed the manuscript.

Competing interests. We, the authors, declare that there are no competing interests present.
Figure C1. Mountain ridges shown in a similar fashion as in previous studies for the Rocky Mountain region. Width of the boxes is given by half the wavelength of excited waves, length is given by the length of the underlying ridge. Boxes have been drawn in order of ascending height so that the highest ridges (which are typically the smaller ones) are visible atop the large scale small amplitude features. Note that the height shown is the height of the single ridge, not the accumulative height of all ridges at that position.

Figure C2. Same as Fig. C1 but for the Southern Andes region, which is considered extensively in this study.

Acknowledgements. The work of SR was funded by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) project UN 311/4-1. In addition, SR and his visit to JB was partly funded by DFG project PR 919/5-1. The work of LK was partly funded by the German
Ministry for Education and Research under grant 01 LG 1907 (project WASCLIM) in the frame of the Role of the Middle Atmosphere in Climate (ROMIC)-program. The work of ME was supported by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) project ER 474/4-2 (MS–GWaves/SV), which is part of the DFG research unit FOR 1898 (MS–GWaves). The work by ME was also supported by the Federal German Ministry for Education and Research (Bundesministerium für Bildung und Forschung, BMBF) project QUBICC (grant no. 01LG1905C), which is part of the Role of the Middle Atmosphere in Climate II (ROMIC-II) program of BMBF.
References


